# A NEW MOTION COMPENSATION APPROACH FOR ERROR RESILIENT VIDEO CODING

*Mengyao Ma*†*, Oscar C. Au*‡*, S.-H. Gary Chan*†

† Dept. of Computer Science, ‡ Dept. of Electrical and Electronic Engineering
Hong Kong University of Science and Technology
{myma,eeau,gchan}@ust.hk

### ABSTRACT

*Multihypothesis motion-compensated prediction* (MHMCP) can be used as an error resilience technique for video coding. Motivated by MHMCP, we propose a new error resilience approach named *Alternative Motion-Compensated Prediction* (AMCP), where two-hypothesis and one-hypothesis predictions are alternatively used with some mechanism. Both theory and simulation results show that in case of one frame loss, the expected converged error using AMCP is smaller than that using two-hypothesis MCP.

## 1. INTRODUCTION

Since the quality of compressed video is vulnerable to errors, video transmission over unreliable Internet is very challenging today. If INTER mode is used in video coding, each frame is predicted from a previously coded frame by *Motion Compensation*, and sent to the decoder. In case of data loss, the corresponding frame will be corrupted, and this error will be propagated to the following frames until the next INTRA-coded frame is correctly received. Due to these facts, it is useful to develop some schemes to improve the *Error Resilience* (ER) capability of the compressed video [1]. Several ER methods have been developed for video communication, such as *Forward Error Correction* (FEC) [2], *Layered Coding* [3], *Multiple Description Coding* (MDC) [4], and *Multihypothesis Motion-Compensated Prediction* (MHMCP) [5]. Our work is motivated by the latter two methods.

In temporal sub-sampling MDC, the video is coded into multiple decodable streams; each with its own prediction and transmission [6]. One simple implementation is odd/even sub-sampling: an even (odd) frame is predicted from the previous even (odd) frame. Since the reference frames are farther in time, the prediction of such approach is not as good as the conventional codec and the compression efficiency is lower. On the other hand, since each stream is separately encoded and transmitted , the corruption of one stream will not affect the other. As a result, the decoder can simply display the correct video stream, or reconstruct the corrupted frame by *Temporal Interpolation*.

In MHMCP, a linear combination of multiple signals (hypothesis) is used to predict each macroblock. Since such prediction is better than that using only one hypothesis, coding efficiency is improved. Besides its coding gain, MHMCP can also improve the error resilience capability of a codec. In [5], each frame (except I frame) is predicted from its previous two frames. The error propagation model at the decoder side is analyzed, which is combined with the encoder predictor to strike a balance between compression efficiency and error resilience capability. In contrary to the odd/even sub-sampling approach, if one frame is corrupted during the transmission, error will propagate to all the following frames and converge at last. We define *Error Ratio* to be the ratio of converged value to the first error, and call this two-hypothesis approach THMCP.

Based on the properties of previous methods, we propose a novel motion compensation approach, where odd/even sub-sampling and two-hypothesis prediction are combined. The novelty is that by using *Alternative Motion-Compensated Prediction* (AMCP), the expected *Error Ratio* at the decoder is less than that obtained from THMCP. The rest of this paper is organized as follows: In Section 2, we describe the working of AMCP and derive its error propagation model. The expected *Error Ratio* of AMCP and THMCP are compared in Section 3. Section 4 shows the simulation results and Section 5 is conclusion.

## 2. ALTERNATIVE MOTION-COMPENSATED PREDICTION (AMCP)

In THMCP [5], frame $\psi(n)$ has two hypotheses and is predicted by

$$\hat{\psi}(n) = h_1\tilde{\psi}(n-1) + h_2\tilde{\psi}(n-2), \tag{1}$$

This work has been supported in part by the Research Grants Council of the Hong Kong Special Administrative Region, China (project no. HKUST6203/02E and HKUST6156/03E), the Innovative Technology Commission (project no. ITS/122/03) and HKUST (project no. DAG04/05.EG34).

0-7803-9134-9/05/$20.00 ©2005 IEEE

where $n \geq 2$ and $h_1 + h_2 = 1$. $\tilde{\psi}(n-k)$ is a motion-compensated prediction from the $k^{th}$ previous reconstructed frame, $k = 1, 2$. Note that if $h_2 = 0$, this becomes a conventional predictor. And if $h_2 = 1$, this is the same as the odd/even sub-sampling method.

Based on the idea of THMCP and odd/even sub-sampling method, we try one simple approach first, where each even frame is predicted from its previous two frames using (1) and each odd frame is predicted from its previous odd frame, as in Figure 1(a). Suppose one frame is lost. If it is an odd frame (with $\circ$), error will propagate to all the following frames; if it is an even frame (with $\times$), the error value will decrease and converge to zero quickly. Motivated by the good characteristics of the latter case, we propose a new ER approach named *Alternative Motion-Compensated Prediction* (AMCP). In AMCP, the video sequence is divided into periodic *Intervals* $\{I_0, I_1, \ldots\}$, which start after an I frame. The frame index within an Interval (*Interval Index*) goes from $0$ to $(2N+1)$, and the $(2N+1)^{th}$ frame is the $0^{th}$ frame of the next *Interval*, as illustrated in Figure 1(b). In each *Interval*, the odd frame is predicted from its previous two frames using (1), and the even frame is predicted from its previous even frame. Here the odd or even frame is defined by its *Interval Index*, instead of its frame index in the video sequence. A special case is $N = 0$, which makes the predictor the same as THMCP.

Consider the case of one frame loss during the transmission. Assume the lost frame is $\psi(l)$ and define error $\epsilon(k)$ to be the difference between the reconstructed $(l+k)^{th}$ frame at the decoder and that at the encoder. Using THMCP, the error propagation model is

$$\epsilon(k) = \frac{1 - (-h_2)^{k+1}}{1 + h_2}\epsilon(0). \tag{2}$$

If the effect of spatial filtering caused by *Sub-Pixel Motion Compensation* is not considered, this error will converge at last with *Error Ratio* $R_1 = \frac{1}{1+h_2}$ [5].

In our AMCP method, without loss of generality, suppose $\psi(l)$ belongs to $I_0$ and the first error is also $\epsilon(0)$. We want to obtain the error propagation model. Similar as the simple approach we state previously, the loss of an odd frame (with $\times$) or an even frame (with $\circ$) will form different error propagations, as in Figure 1(c).

**Case1** A frame with an odd *Interval Index* is lost and the error propagation within $I_0$ is

$$\begin{array}{ll} \epsilon_1(2n) = h_2^n \epsilon(0), & n \geq 1, \\ \epsilon_1(2n+1) = 0, & n \geq 0. \end{array} \tag{3}$$

**Case2** A frame with an even *Interval Index* is lost and the error propagation within $I_0$ is

$$\begin{array}{ll} \epsilon_2(2n) = \epsilon(0), & n \geq 1, \\ \epsilon_2(2n+1) = (1 - h_2^{n+1})\epsilon(0), & n \geq 0. \end{array} \tag{4}$$

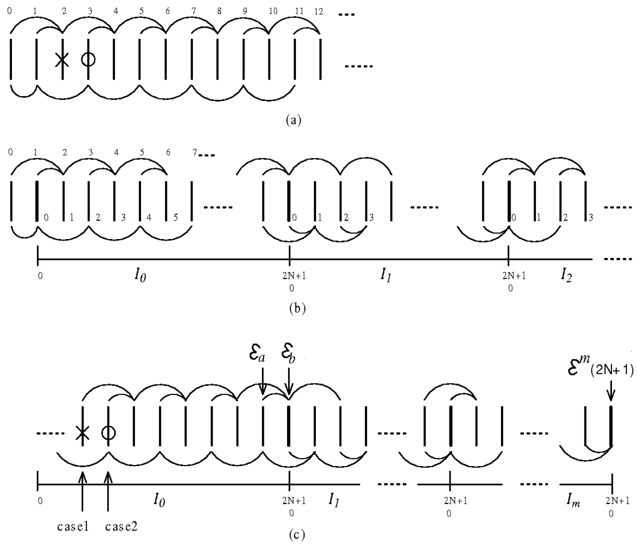


**Fig. 1**. Illustration of AMCP and its error propagation.

$\epsilon_1(k)$ ($\epsilon_2(k)$) is the error at the $(l+k)^{th}$ frame. To analyze the error propagation in the following *Intervals*, define the error in the $r^{th}$ frame of *Interval m* is $\varepsilon^m(r)$ ($r \in [0, 2N+1]$), and $\varepsilon^{m+1}(0) = \varepsilon^m(2N+1)$. For simplicity, suppose the errors of the last two frames in $I_0$ are known: $\varepsilon^0(2N) = \varepsilon_a$ and $\varepsilon^0(2N+1) = \varepsilon_b$. Using similar deriving process as (3) and (4), we can obtain

$$\varepsilon^m(2N+1) = \frac{\beta + (-\beta)^{m+1}}{1+\beta}\varepsilon_a + \frac{1 - (-\beta)^{m+1}}{1+\beta}\varepsilon_b, \tag{5}$$

where $\beta = h_2^{N+1}$. By (3) and (4), we can see that if the frame loss is case1, the values of $\varepsilon_a$ and $\varepsilon_b$ will be much smaller than those of case2, thus leading to a smaller converged value of $\varepsilon^m(2N+1)$.

## 3. EXPECTED ERROR RATIO OF AMCP

From the analysis of the previous section, we can see that the values of $\varepsilon_a$ and $\varepsilon_b$ are determined by the position of the first error, using (3) for case1 or (4) for case2. Consider the condition of one frame loss in $I_0$. Its *Interval Index* can be $i = 0, 1, \ldots, 2N$, each with equal probability. $(2N+1)$ is not included since it can be counted as the first one in the next *Interval*. The expected value of $\varepsilon_a$ and $\varepsilon_b$ can be calculated as

$$E[\varepsilon_a] = \frac{N+1}{2N+1}\epsilon(0), \tag{6}$$

$$E[\varepsilon_b] = \frac{N+1-h_2^{N+1}}{2N+1}\epsilon(0). \tag{7}$$

Combine (6), (7) with (5), we can obtain the expected converged error of AMCP. The expected *Error Ratio* is

$$R_2 = \frac{N + N\beta + 1}{(1+\beta)(2N+1)}, \tag{8}$$

where $\beta = h_2^{N+1}$ ,$h_2 < 1$. Define $\Delta = R_2 - R_1$. When $N$ is fixed and $h_2 \in (0,1)$, $\Delta$ is a monotone increasing function of $h_2$. Its range is $(-\frac{N}{2N+1}, 0)$.

Based on the previous analysis, we can make the following conclusions, which will be verified in the next section:

- When $0 < h_2 < 1$, $R_2 < R_1$, for fixed N. In other words, the expected converged error using AMCP is smaller than that using THMCP.
- When $h_2$ approaches 0, AMCP performs much better than THMCP. On the other hand, a small value of $h_2$ makes both $R_2$ and $R_1$ large. If $h_2 = 0$, THMCP becomes a conventional codec and $\epsilon(0)$ will propagate to all the following frames. Similar result for case2 of AMCP. If the loss is case1, only one frame is corrupted and the rest of the stream remains correct!
- When $h_2$ approaches 1, AMCP and THMCP perform similar, both with decreasing *Error Ratio*. A special case is $h_2 = 1$, both methods converge to the odd/even sub-sampling method, thus *Error Ratio* cannot converge.

## 4. SIMULATION RESULTS

We use the JVT reference software version 8.2 for our simulations [7]. The first 200 frames of video sequence *Foreman* (QCIF) are used for testing, encoded at 30fps and only the first frame is I frame. B frame is used to implement the frame with two references. Fixed QP is used, 28 for I frame and 30 for both P and B frames. In order to analyze error propagation, Intra-MB is not used in P and B frames. We simulate the case of one frame loss, and the lost frame is error concealed by copying the previous frame.

To test the error propagation of AMCP, we randomly select one frame to be lost for case1 (Figure 2(a,b)) and another one for case2 (Figure 2(c,d)). PSNR at the decoder after that frame is plotted, which is defined to be 99.99 for no error. THMCP, AMCP with $N = 2$ and $N = 5$ are tested for each case, with weight $h_2 = 0.1$ or $h_2 = 0.9$. The first distortions (PSNR) for these three methods are very close: around 23.7 dB for case1 and 24.8 dB for case2. From the figure we can see that for a large $h_2$, AMCP and THMCP perform similar for both cases, except that AMCP converges slower than THMCP. For a small $h_2$, AMCP performs much better than THMCP in case1 and only a little worse in case2. A smaller $N$ makes AMCP more close to THMCP.

The expected distortion for one frame loss is also compared. The *Mean Square Error* (MSE) at the decoder is used as the distortion measure. For AMCP, the frames in $I_1$ with *Interval Index* $0$ to $2N$ are selected for loss, one at a time. For each of the loss, the MSE at the decoder is computed, and the average of these $(2N + 1)$ MSEs is used as the expected distortion. Figure 3(a) and (b) are the simulation results for $N = 5$, where the expected distortion is translated into PSNR for representation. The MSEs for the same losses in THMCP are also obtained, and the identical calculation as AMCP is used to get the expected distortion. From the figure we can see that with $h_2 = 0.9$, AMCP and THMCP perform similar, while with $h_2 = 0.1$, the expected distortion of AMCP is much smaller than that of THMCP.

## 5. CONCLUSION

In this paper, we propose an error resilience approach named *Alternative Motion-Compensated Prediction* (AMCP), which can be a generalization of two-hypothesis MCP. We prove that the expected converged error using AMCP is smaller than that using THMCP, in case of one frame loss. Simulation results are given for the justification.

## 6. REFERENCES


[1] Y. Wang, S. Wenger, J. Wen, and A. K. Katsaggelos. (2000, July) Review of error resilient coding techniques for real-time video communications. SPM_W3K2.pdf. [Online]. Available: http://vision.poly.edu:8080/paper/

[2] Y. Mei, W. Lynch, and L. N. Tho, "Joint forward error correction and error concealment for compressed video," in *Proc. ITCC*, Apr. 2002, pp. 410 – 415.

[3] R. Aravind, M. Civanlar, and A. Reibman, "Packet loss resilience of MPEG-2 scalable video coding algorithms," in *IEEE Trans. Circuits Syst. Video Technol*, vol. 6, Oct. 1996, pp. 426 – 435.

[4] M. Orchard, Y. Wang, V. Vaishampayan, and A. Reibman, "Redundancy rate-distortion analysis of multiple description coding using pairwise correlating transforms," in *Proc. IEEE ICIP*, Oct. 1997, pp. 26–29.

[5] S. Lin and Y. Wang, "Error resilience property of multihypothesis motion-compensated prediction," in *Proc. IEEE ICIP*, June 2002, pp. 545 – 548.

[6] J. Apostolopoulos, "Reliable video communication over lossy packet networks using multiple state encoding and path diversity," in *Proc. SPIE VCIP*, Jan. 2001, pp. 392 – 409.

[7] Jvt reference software, version 8.2. [Online]. Available: http://iphome.hhi.de/suehring/tml/download/

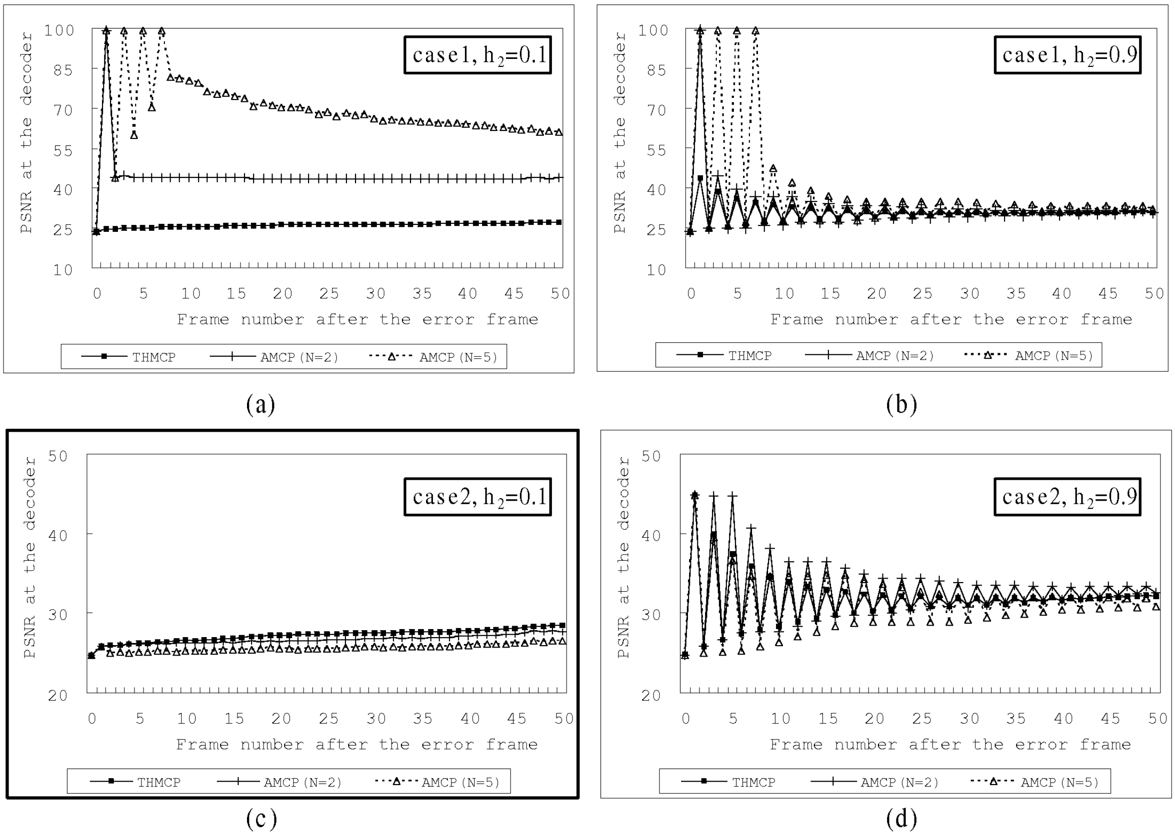


**Fig. 2**. Error propagation at the decoder for one frame loss: (a,b) for case1 and (c,d) for case2.

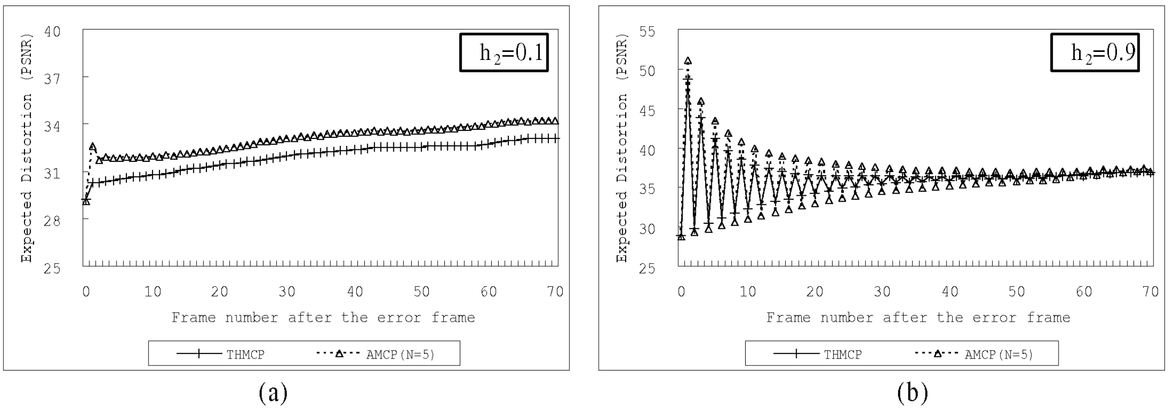


**Fig. 3**. Expected distortion at the decoder for one frame loss.